Workpackage 3

Link: www.fcg-nwt.org

I. Objectives

In the previous report, it was explained that investigating the analogy between disease-associated gene interactions and language evolution (1) requires a more profound understanding of regulatory mechanisms in the immune system as language processing or coding mechanisms and (2) requires that the role and emergence of compositionality and grammar in coding are better understood. Several approaches were proposed to tackle these issues further. The approach based on Artificial Chemistry looked especially promising. The results of this line of research will be discussed in Sections II.1 and II.2. Section II.1 mainly is about advancements in theoretical semiotic dynamics, while Section II.2 shows how the developed theory can be applied both to natural language and the immune system. The other line of research concerns the further development of Fluid Construction Grammar as a software tool for performing experiments involving grammar. This will be discussed in section II.3

II. Progress towards objectives – tasks worked on and achievements made with reference to planned objectives, identify contractors involved

II.1 Progress towards the development of a general theory of conventionalization

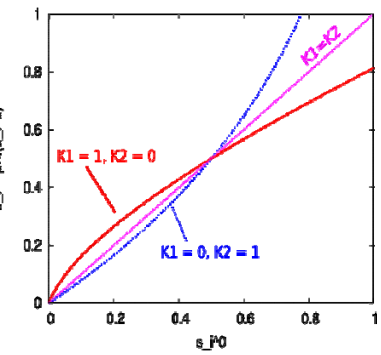
In order to find an analogy between coding in language and in the immune system, we developed a *general theory of conventionalization*. Conventionalization is the process by which conventions arise in a population of code users through self-organization. Language is the product of conventionalization mechanisms itself. As a code it is also used to negotiate about other conventions. Similarly, the genetic code is a fossil that reflects the dynamic, primordial conditions in the RNA world [1,2]. Currently it is used to distinguish between *self* and *non-self* in the immune system. This is a conventionalization problem as well. For example, vaccination and immune therapy induce conventionalization mechanisms that re-configure it.

The theory we developed is general in the sense that it treats meanings and signs as abstract entities in their own right, with a similar status as for example rigid bodies in Theoretical Physics. In this sense, conventionalization can be studied independent of a particular domain. In particular, code users are modeled as chemical reactor tanks interacting with their environment through the absorption and secretion of artificial chemical substances *representing* meanings and signs. This has the additional advantage that code users become mathematically well defined objects. Furthermore, the coding and learning behavior of a code user is then entirely determined by the Artificial Chemistry governing the chemical reactions taking place in the reactor tank. This means that we have Chemical Reaction Network Theory (CRNT, [3]) and Chemical Organization Theory [4] at our disposal for investigating this behavior. And since the macro behavior of a population of adaptive is influenced by the behavior of individual code users and vice versa, these tools can then also be employed study

conventionalization in general. In particular, this can be achieved by subjecting the behavior of agents to a *response analysis* [5].

In order to assess the adequacy of this approach, we have formalized the notion of *Injective Organic Codes (IOCs)* based on a general definition of organic codes in [6] and an earlier formalization for *Binary Injective Organic Codes* in [7]. Using the tools described, we were able to characterize the desired behavior of code users capable of evolving and using arbitrary *IOCs* (see Figures 1 and 2). In order to assess the relevancy of the approach in the context of language, we then identified the notion of *IOCs* with two benchmark problems well known in Semiotic Dynamics and Evolutionary Linguistics, namely the *naming* and *guessing games* [8,9]. We also identified a concrete code user model within the context of the immune system. These findings expose the immune system as a distributed information processing and decision making coding system. This will be further discussed in the following section.

Figure 1. *Response analysis (top) of an Artificial Chemistry (bottom) governing the behavior of a coding agent. In the reactions, 'm', 's', and 'c' stand for meaning, sign and codemaker respectively (see also [6,7,10,11]). K1 and K2 are the reaction rates of the last two reactions in the Artificial Chemistry. The X-axis shows the population usage frequency of signs, relative to a reference frequency of 0.5. The Y-axis shows the resulting usage frequency of an individual code user observing the population behavior. If K1>K2, then signs that are used less than average in the population will be preferred by the individual, and vice versa if K1<K2. This means that a positive feedback loop driving all code users to use the same, conventional sign will only be possible if K1 < K2.*



$$m + c_j \xrightarrow{k_s} s_j$$

$$m + s_j \xrightarrow{k_c} c_j$$

$$m + s_j + c_j \xrightarrow{k_a} 2s_j,$$

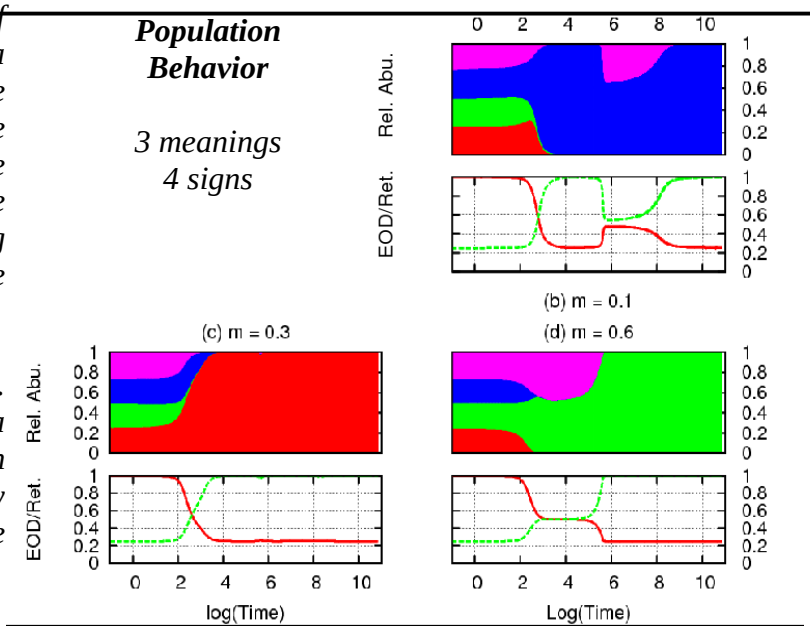
$$m + s_j + c_{j'} \xrightarrow{\kappa_1} m + s_{j'} + c_{j'}$$

$$m + s_j + c_{j'} \xrightarrow{\kappa_2} m + s_j + c_j$$

The graph can be obtained directly from the Artificial Chemistry by using the reactor tank metaphor for code users, illustrating the strength of the approach both for constructing novel agent behaviors and for assessing existing ones [5].

Figure 2. *Behavior of a population of code users self-organizing a conventional Injective Organic Code (IOC). Three meanings need to be expressed. Each graph shows the temporal evolution of the relative (regulated) affinity of each meaning for all signs. Different colors denote the affinity for different signs.*

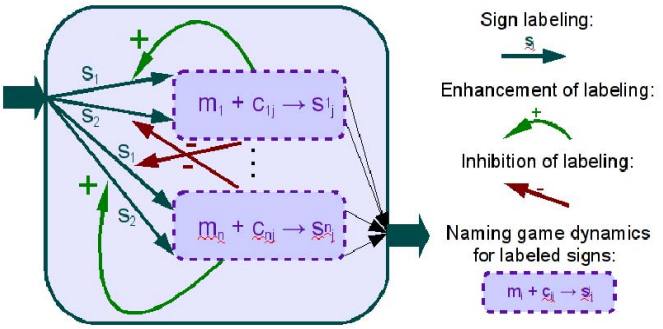
Initially, four signs are available. Eventually, each meaning selects a different sign and one redundant sign is discarded. The bottom graphs show the corresponding evolution of the population coherence [10,11].



II.2 Progress towards developing an analogy between language and the immune system

Figure 3 shows a schematic description of the Artificial Chemistry governing the behavior of the code users in the population of Figure 2. Figure 4 schematically depicts a model for T-helper cell differentiation in the immune system. This model, which was taken from [12], is a simplification, and recognized as such by biomedical researchers. For example, on the cellular level, other Th cell subsets (Th3) also regulate the response by playing an inhibitory role [13]. However, the Th1/Th2 model is consistent with a large body of experimental findings and is generally considered to capture a core regulatory mechanism of the immune system. As the figure illustrates, this mechanism is essentially a conventionalization mechanisms by which a system of distributed code users can reach a global, coherent language or decision through local interactions. In particular, it allows T-helper cells, communicating through the secretion and absorption of cytokines, like IFN-gamma and IL-4, to achieve a coherent and targeted immune response for a number of different threats. This result was a joint achievement obtained through extensive discussions with partners 1 (UGOT) and 2 (RR-HF). It was submitted to the Journal of Biosemiotics together with an outline of the theory that was discussed in Section II.1 and is currently still under review.

Figure 3. *Schematic representation of the Artificial Chemistry governing the behavior of code users in the population of Figure 2. Each smaller box within the large box introduces a naming game reaction network (see Figure 1): one for each available meaning. Labeling mechanisms decouple the different naming games dynamics. Self-enhancing and cross-inhibitory mechanisms re-couple them in such a way that a unique* IOC *will become selected.*



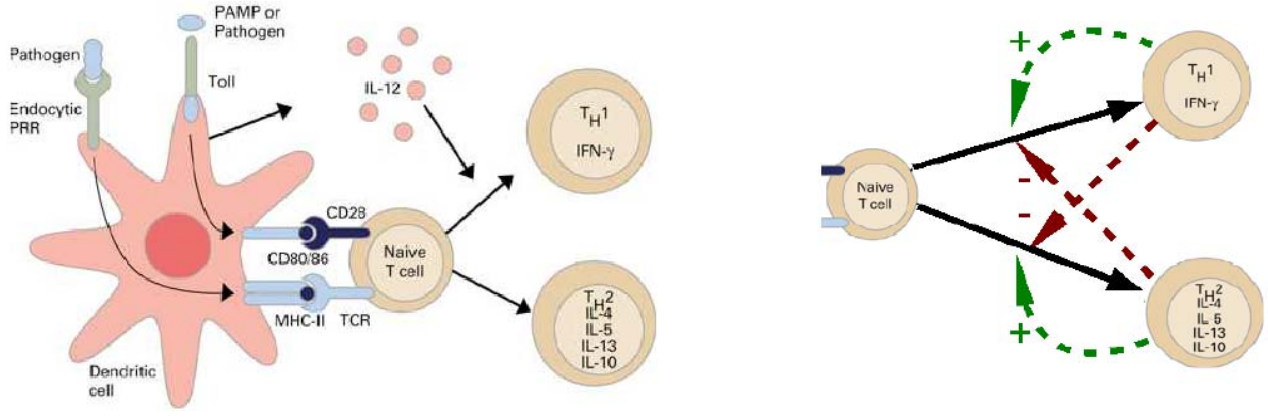


Figure 4. Left: *Th1/Th2 T-helper cell differentiation in the immune system involving a dendritic antigen presenting cell (APC) and several cytokines [12].* ***Right:*** *Schematic representation of this regulatory mechanism. Th1 cells promote the formation of more Th1 cells. They inhibit the formation of Th2 cells. Th2 cells promote Th2 cell formation and inhibit Th1 cell formation. Schematically, this is the same mechanisms as the one used by the code users in Figures 2 and 3 for self-organizing an* IOC.

II.3 Progress in understanding grammar and the development of simulation software

With the developments as discussed in sections II.1 and II.2, we have been able to formulate a formal and concrete analogy between language and the immune system. Nevertheless, *IOC*s are still much simpler than the regulatory networks governing natural language and the complete immune system. As was already explained in the previous report, in order to make further progress we will also need to understand the function of grammar in language, and of larger regulatory networks in the immune system. For this, we have been developing Fluid Construction Grammar (FCG), as well as performed a number of case studies concerning the automatic parsing and production of natural language. This is illustrated by a number of publications (see [14-17]) and by a regular sequence of software releases through http://www.fcg-net.org/.

III. Deviations from the project work program, and corrective actions taken/suggested: identify the nature and the reason for the problem, identify contractors involved

In the previous report, it was already explained why it was found infeasible to start with data obtained from real patients. The major obstacles and corrective actions that were proposed are summarized again in Section I. As was discussed in Section II., these actions indeed lead to substantial progress: there have been several relevant publications (see references) as well as a forthcoming joint vision paper that further advances the field of Biosemiotics as a formal, scientific theory. Also taking into account the regular FCG releases, this means that we were able to reach all of the milestones and deliverables, except for the usage of data obtained from real patients. The main obstacle further prohibiting this is the relative simplicity of the languages that can be tackled so far (i.e. Injective Organic Codes) compared to the complexity of the grammatical and regulatory mechanisms governing natural language and the immune system. These issues are currently still further investigated, also outside the context of this Project. The results that were obtained within the context of this project therefore form a valuable contribution to the field of semiotic dynamics and immunology in general.

IV. List of deliverables, including due and actual/foreseen submission date

D10: Software that performs computer simulations. Due month 24. Regular releases of the Babel2 simulation software platform that includes Fluid Construction Grammar were (and continue to be) released through http://www.fcg-net.org. This software is being used by a growing number of researchers.

D11: Comprehensive overview of how simulation results can be analyzed. This mainly forms the subject of the paper [11], which is under revision and will be published soon.

V. List of milestones, including due and actual/foreseen achievement date

M7 Investigation of analogy between disease-related gene-interactions and language evolution (year 1) , M8 Case study (year 2) and M9 Analysis of obtained results (year 3). A core regulatory mechanism in the immune system was identified as a conventionalization mechanism that is capable of self-organizing arbitrary Injective Organic Codes in a population of distributed code users. It is the same mechanism that solves the guessing game problem from Evolutionary Linguistics (see [11]).

V. Selected References

(names of people from the VUB AI-lab or otherwise directly involved in the project are in **bold**)

[1] Di Giulio, M. The late stage of genetic code structuring took place at a high temperature,.Gene 261p.198-195 (2000)

[2] Jestin, A.L. and Kempf, A., Optimization models and the structure of the genetic code. J.Mol.Evol 69p.452-457 (2009)

[3] Craciun, G. and Feinberg, M. Multiple Equilibria in Complex Chemical Reaction Networks: Extensions to Entrapped Species Models. Systems Biology 153(4)p.179-186 (2006)},

[4] Dittrich, P. and di Fenizio, P.S. Chemical Organization Theory. Bull. Math. Biol. 69(4) (2007)

[5] **De Vylder, B.**, The Evolution of Conventions in Multi-Agent Systems, unpublished doctoral d issertation, VUB Artificial Intelligence Lab (2007)

[6] Barbieri, M. Biosemiotics: a new understanding of life. Naturwissenschaften 95(7)p.577-59 (2008)

[7] Dennis Gorlich and Peter Dittrich, Identifying Molecular Organic Codes in Reaction Networks. In: Campis, G. and Szathmary, E. (eds.) Proceedings of the 10th European Conference on Artificial Life, LNCS/LNAI 5777, (2009)

[8] **Steels, L.** and McIntyre, A. Spatially Distributed Naming Games. Adv.in Compl. syst. (4) (1999)

[9] Andrea Baronchelli, Vittorio Loreto and **Luc Steels**: In-depth analysis of the Naming Game dynamics: the homogeneous mixing case. In: Int. J. of Mod. Phys. C **19**, 785 (2008)

[10] **De Beule, J.** A Chemical Model of the Naming Game. In: Campis, G. and Szathmary, E. (eds.) Proceedings of the 10th European Conference on Artificial Life, LNCS/LNAI 5777, (2009),

[11] **De Beule**, **J., Eivind, H.** and **Benson, M.** Introducing dynamics into the field of Biosemiotics with examples from language and immunology. Submitted.

[12] Mendoza, L. A network model for the control of the differentiation process in Th-cells. Biosystems, 84(2)p101-114 (2006)

[13] **Benson, M.** and Adner, M. and Cardell, L.O. Cytokines and cytokine receptors in allergic rhinitis: how do they relate to the Th2 hypothesis in allergy? Clin Exp Allergy 31p.361-367 (2001)

[14] Wellens, P. and **De Beule, J.** Priming through Constructional Dependencies, a case study in Fluid Construction Grammar. Proceedings of the 8th evolution of Language Conf.

[15] Micelli, V., van Trijp, R. and **De Beule, J.** (2009). Framing Fluid Construction Grammar. In: Taatgen, N.A. and van Rijn, H. (eds.) *Proceedings of the 31th Annual Conference of the Cognitive Science Society*. p. 3023--3027. Cognitive Science Society.

[16] Gerasymova, K., **Steels, L.** and van Trijp, R. (2009). Aspectual Morphology of Russion Verbs in Fluid Construction Grammar. Ibid.p. 1370--1375.

[17] **De Beule, J.** and Micelli, V. Who Framed Fluid Construction Grammar? In: Boas, H.(ed.) Computational Approaches to Construction Grammar and Frame Semantics (2010)